Q1

a) Because x, y are independent

$$p(\mu,\pi|x,y) = p(\mu|x)p(\pi|y)$$

According to bayes rule

$$p(\mu|x) = \frac{p(x|\mu)p(\mu)}{\int p(x|\mu)p(\mu)d\mu}$$

$$= \frac{e^{-\frac{\lambda(x-\mu)^2}{2}} e^{-\frac{\gamma\mu^2}{2}}}{\left(\frac{\lambda+\gamma}{2\pi}\right)^{-\frac{1}{2}}}$$

$$= \left(\frac{\lambda+\gamma}{2\pi}\right)^{\frac{1}{2}} e^{-\frac{(\lambda+\gamma)\left(\mu-\frac{\lambda x}{\lambda+\gamma}\right)^2}{2}} = \mathrm{Normal}\left(\frac{\lambda x}{\lambda+\gamma}, (\lambda+\gamma)^{-1}\right)$$

For $p(\pi|y)$

$$p(\pi|y) = \frac{p(y|\pi)p(\pi)}{\int p(y|\pi)p(\pi)d\pi}$$

$$= \frac{1}{B(\alpha')}\prod_{k=1}^{3} \pi(k)^{\alpha_k - 1 + 1(y=k)} = \mathrm{dirichlet}(\alpha')$$

$$\alpha' = \{\alpha_k + 1(y=k)\}$$

Therefore

$$p(\mu,\pi|x,y) = \left(\frac{\lambda+\gamma}{2\pi}\right)^{\frac{1}{2}} e^{-\frac{(\lambda+\gamma)\left(\mu-\frac{\lambda x}{\lambda+\gamma}\right)^2}{2}} * \frac{1}{B(\alpha')}\prod_{k=1}^{3} \pi(k)^{\alpha_i - 1 + 1(y=k)}$$

b) According to the note and independence

$$p(x_2, y_2|x_1, y_1) = \int p(x_2|\mu)p(\mu|x_1)d\mu * \int p(y_2|\pi)p(\pi|y_1)d\pi$$

by using the outcome from part a, we can calculate the integral

$$\int p(x_2|\mu)p(\mu|x_1)d\mu = \left(\frac{\lambda(\lambda+\gamma)}{2\pi}\right)^{\frac{1}{2}} \exp\{-\frac{\lambda}{2}x_2^2 + \frac{\lambda^2(x_1+x_2)^2}{2(2\lambda+\gamma)} - \frac{\lambda^2 x_1^2}{2(\lambda+\gamma)}\}$$

$$\int p(y_2|\pi)p(\pi|y_1)d\pi = \frac{\prod_{k=1}^{3}(\alpha_k + 1(y_1 = k))^{1(y_2=k)}}{\sum_{i=1}^{3}(\alpha_i) + 1}$$

Therefore

$$p(x_2, y_2|x_1, y_1) = \left(\frac{\lambda(\lambda+\gamma)}{2\pi}\right)^{\frac{1}{2}} \exp\{-\frac{\lambda}{2}x_2^2 + \frac{\lambda^2(x_1+x_2)^2}{2(2\lambda+\gamma)} - \frac{\lambda^2 x_1^2}{2(\lambda+\gamma)}\}\frac{\prod_{k=1}^{3}(\alpha_k + 1(y_1 = k))^{1(y_2=k)}}{\sum_{i=1}^{3}(\alpha_i) + 1}$$

Q2

According to the note

$$q(c) = p(c|\lambda, x)$$

$$= \frac{p(x|\lambda, c)p(c)}{\int p(x|\lambda, c)p(c)dc}$$

$$\propto e^{-\lambda_c}\frac{\lambda_c^x}{x!}\theta_c$$

$$\lambda_i = \arg\max_{\lambda_i} L(\lambda) = x$$

EM algorithm

1. Initializing λ_0 in some way
2. For iteration $t = 1,2, \ldots, T$

   (a)E-step: Calculate $E = E_{q_t}[c]$, where

$$E_{q_t}[c] = \sum_{i=1}^{k} e^{-\lambda_{i(t-1)}} \frac{\lambda_{i(t-1)}^{x}}{x!} \theta_{i(t-1)}\, i$$

   (b)M-step: Update vector λ using the expectations above in the following equations

$$\lambda_i = x$$

   (c)Calculate $\ln p(x|\lambda)$ using the equation

$$\ln p(x|\lambda_t) = \ln e^{-\lambda_t} \frac{\lambda_{it}^{x}}{x!}$$

   Here λ_t is one dimension of λ_t vector, since every dimension is the same

Q3

According to the note:

$$q(\theta) \propto \exp\{E_{q(\lambda_1,\ldots,\lambda_n)}[\ln p(x|\lambda) + \ln p(\lambda|\theta) + \ln p(\theta)]\}$$

$$\propto \exp\{\sum_{i=1}^{n} E_q[\ln p(\lambda_i|\theta)]\} p(\theta)$$

$$\propto \theta^{na+b-1} e^{-(\sum_{i=1}^{n} E_q[\lambda_i]+c)\theta} = Gamma(b', c')$$

$$b' = na + b \;,\quad c' = \sum_{i=1}^{n} E_q[\lambda_i] + c$$

$$q(\lambda_i) \propto \exp\{E_{q(\theta)\prod_{j\neq i} q(\lambda_j)}[\ln p(x|\lambda) + \ln p(\lambda|\theta) + \ln p(\theta)]\}$$

$$\propto \exp\{\sum_{i=1}^{n} E_q[\ln p(\lambda_i|\theta)] + E_q[\ln p(x_i|\lambda)]\}$$

$$\propto \lambda_i^{x_i+a} e^{-(E_q[\theta]+1)\lambda_i} = Gamma(a', \theta')$$

$$a' = a + x_i \;,\quad \theta' = E_q[\theta] + 1$$

Inputs: Data and definitions $q(\theta) = Gamma(b', c')$ and $q(\lambda_i) = Gamma(a', \theta')$

Outputs: Values for b', c', a', θ'

1. initialize $b'_0, c'_0, a'_0, \theta'_0$ in some way
2. For iteration $t = 1,\ldots,T$

   Update $q(\theta)$ by setting

$$b' = na + b, \quad c' = \sum_{i=1}^{n} \frac{a'_{i(t-1)}}{\theta'_{t-1}} + c$$

   Update $q(\lambda_i)$ by setting

$$a' = a + x_i, \quad \theta' = \frac{b'_{t-1}}{c'_{t-1}} + 1$$

Evaluate $\mathcal{L}(b_t', c_t', a_t', \theta_t')$ to access convergence

$$\mathrm{L}(b_t', c_t', a_t', \theta_t') = E_q[\ln(x, \lambda, \theta)] - E_q[lnq(\lambda)] - E_q[lnq(\theta)]$$

$$= \sum_{\mathrm{i}=1}^{\mathrm{n}} (x_i(\psi(a_i') - ln\theta') - lnx_i! - \frac{a_i'}{\theta'}) + na(\psi(b') - lnc') - nln\Gamma(a) - n(a-1)ln\theta'$$

$$+ \sum_{i=1}^{n} ((a-1)\psi(a_i') - \frac{a_i' b'}{\theta' c'}) + blnc - ln\Gamma(b) + (b-1)(\psi(b') - lnc') - \frac{cb'}{c}$$

$$- nln\theta' + \sum_{i=1}^{n} (a_i' + ln\Gamma(\, a_i{}') + (1 - a_i{}')\psi(\mathrm{a}_\mathrm{i}')) + b' - lnc' + ln\Gamma(b')$$

$$+ (1 - b')\psi(b')$$